Yield gap decomposition to identify constraints to crop production in farmers’ fields

João Vasco Silva (PhD)

j.silva@cgiar.org

Agronomy-at-scale Data Scientist

CIMMYT-Zimbabwe

Addis Ababa, May 22nd 2023

CIMMYT

Table of contents

1. Introduction
2. Concepts of production ecology
3. What is yield gap decomposition?
4. Processed-based crop models
5. Data-driven approaches
    1. Boundary line analysis
    2. Stochastic frontier analysis
    3. Machine learning
6. Summary
7. Strategic applications

CIMMYT

Introduction

1. Yield gaps to inform food security assessments and delineate scope for sustainable intensification.

2. Understand the relative contribution of growth defining, limiting, and reducing factors to actual farm yields.

3. Increasing availability of farmer field data, and environmental data, both across space and over time.

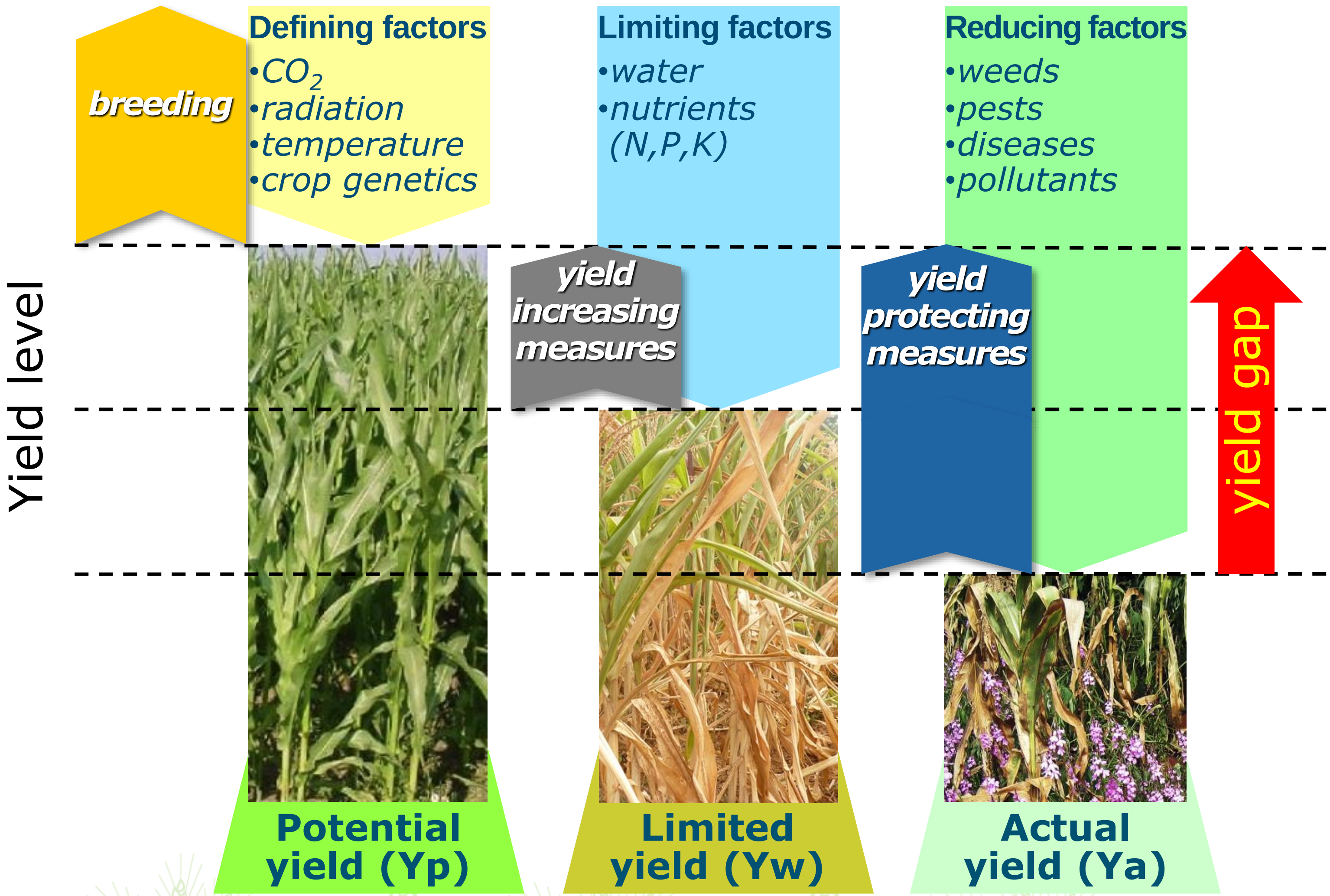
4. Prioritization of R&D based on the most limiting factors to crop production.

Objective: Overview of modelling and data-driven and approaches** for yield gap decomposition using farmer field data.

***Expert-based or experimental assessments not covered.*

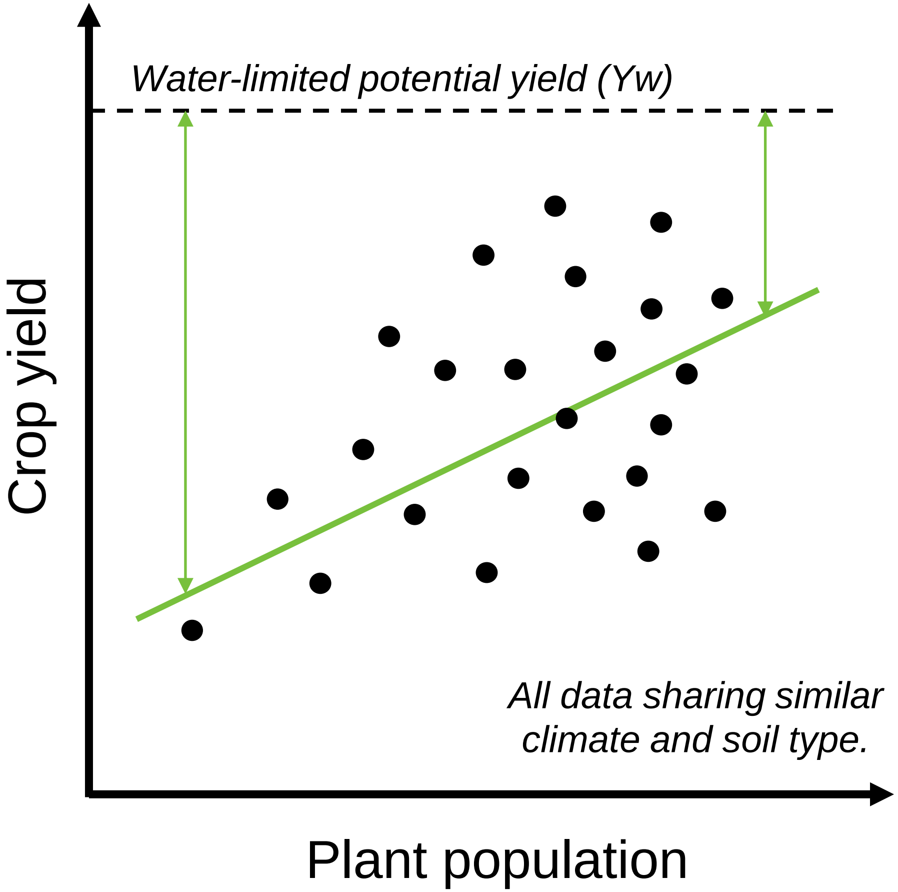
CIMMYT MR

Concepts of production ecology



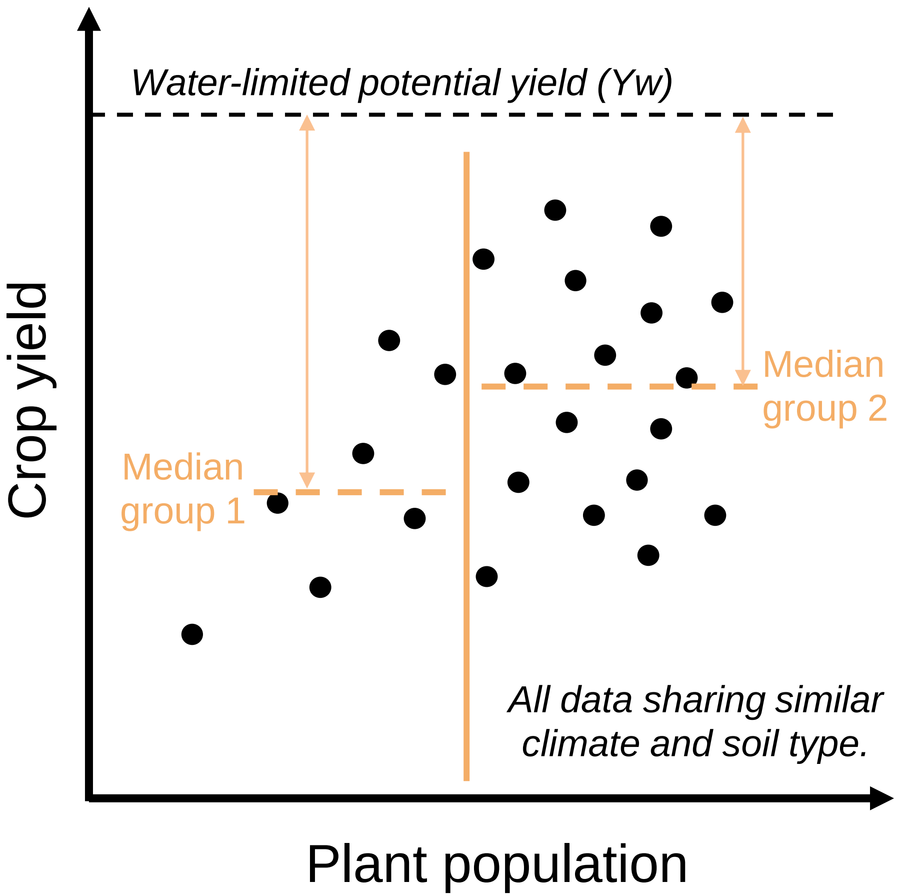
van Ittersum and Rabbinge (1997, FCR)

What is yield gap decomposition?



Regression approach: $y = ax + b + \varepsilon$
Crop yield increases with plant population, on average. Slope indicates rate of change.

What is yield gap decomposition?

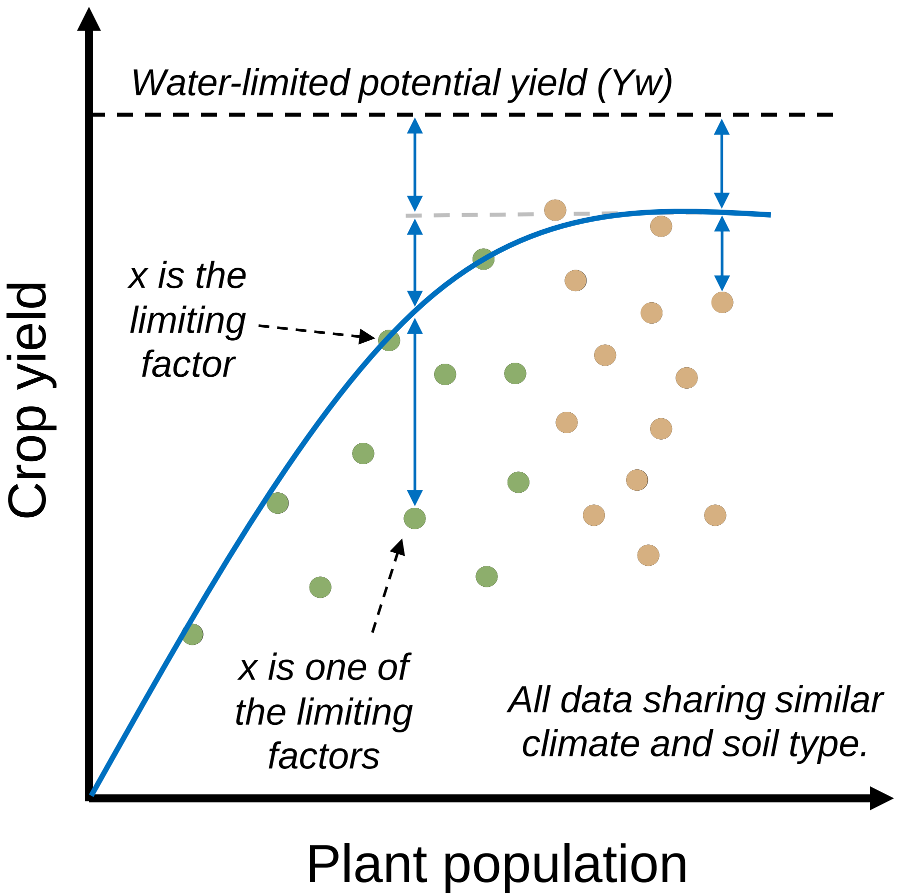


Regression approach: $y = ax + b + \varepsilon$
Crop yield increases with plant population, on average. Slope indicates rate of change.

Tree-based approach: *non-parametric*
Data space partitioned into a group with low plant population and low yield, and a group with high plant population and high yield.

CIMMYT

What is yield gap decomposition?



Regression approach: $y = ax + b + \varepsilon$
Crop yield increases with plant population, on average. Slope indicates rate of change.

Tree-based approach: *non-parametric*
Data space partitioned into a group with low plant population and low yield, and a group with high plant population and high yield.

Frontier approach:
In half the observations the yield gap is due to plant population. In the other half, the yield gap is explained by factors other than plant population.

Yield gap analysis is related to analysis of actual yield variability, but it is not the same!

Process-based crop models

Simulation of yield levels


Field Crops Research 296 (2023) 108891

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Catching-up with genetic progress: Simulation of potential production for modern wheat cultivars in the Netherlands

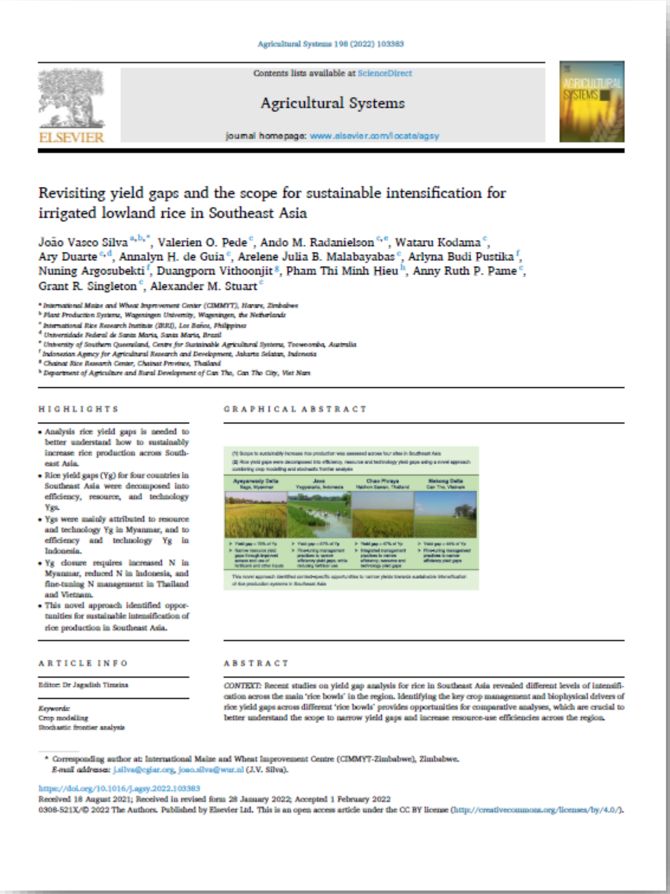
H.N.C. Berghuijs, J.V. Silva, H.C.A. Rijk, M.K. van Ittersum, F.K. van Evert, P. Reidsma

ARTICLE INFO

ABSTRACT

1. Introduction

Simulation of mgt practices


Agricultural Systems 198 (2022) 103383

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

Revisiting yield gaps and the scope for sustainable intensification for irrigated lowland rice in Southeast Asia

João Vasco Silva, Valerien O. Pede, Ando M. Radanielson, Wataru Kodama, Ary Duarte, Annalyn H. de Guia, Arelene Julia B. Malabayabas, Arlyna Budi Pustika, Nuning Argosubekti, Duangporn Vithoonjit, Pham Thi Minh Hieu, Anny Ruth P. Pame, Grant R. Singleton, Alexander M. Stuart

HIGHLIGHTS

GRAPHICAL ABSTRACT

ARTICLE INFO

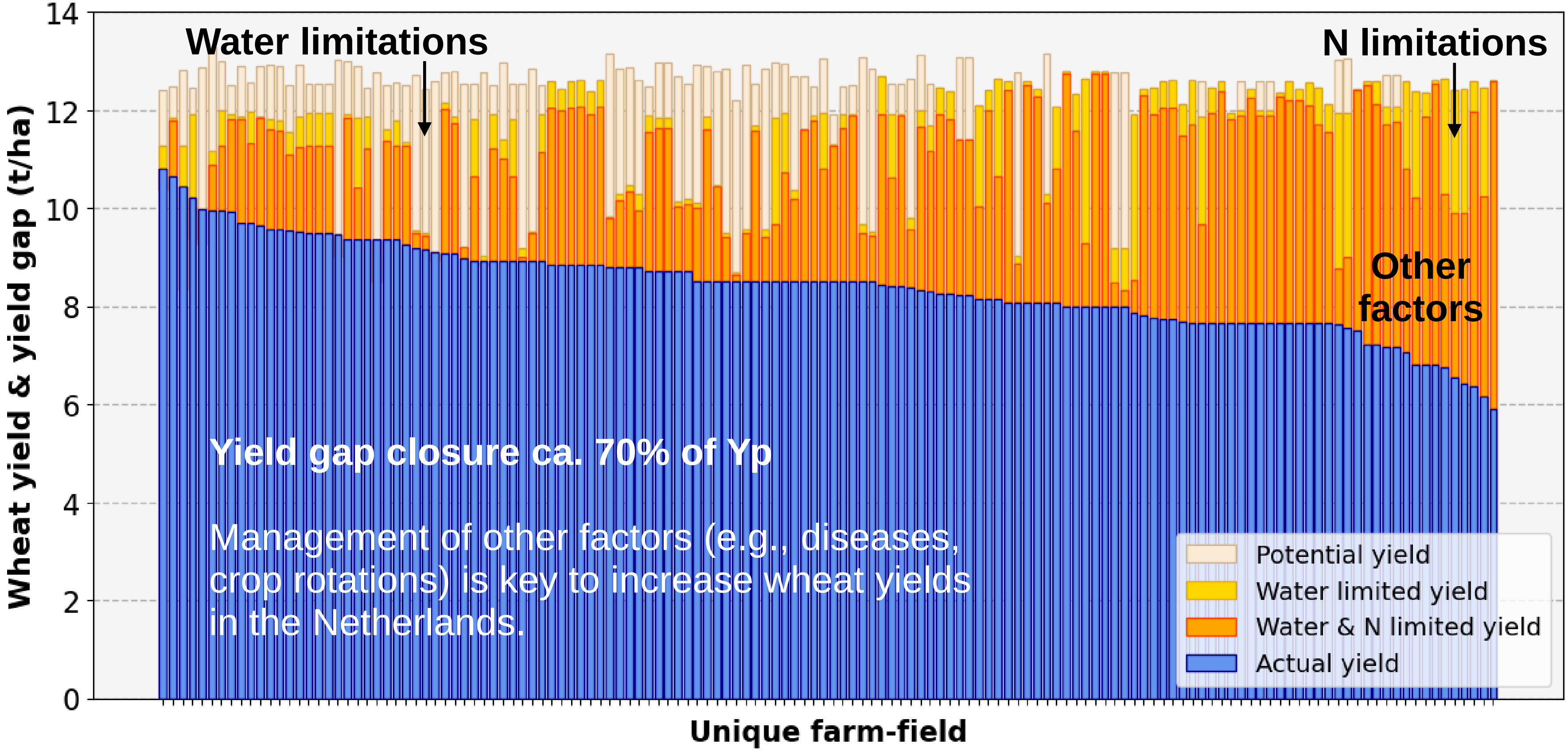
ABSTRACT

Simulation of yield levels

- **Winter wheat crops in the Netherlands**:
  - High-yielding crop, due to high input use and intensive management
  - Important in the rotation to ensure high-yields of tuber/root crops

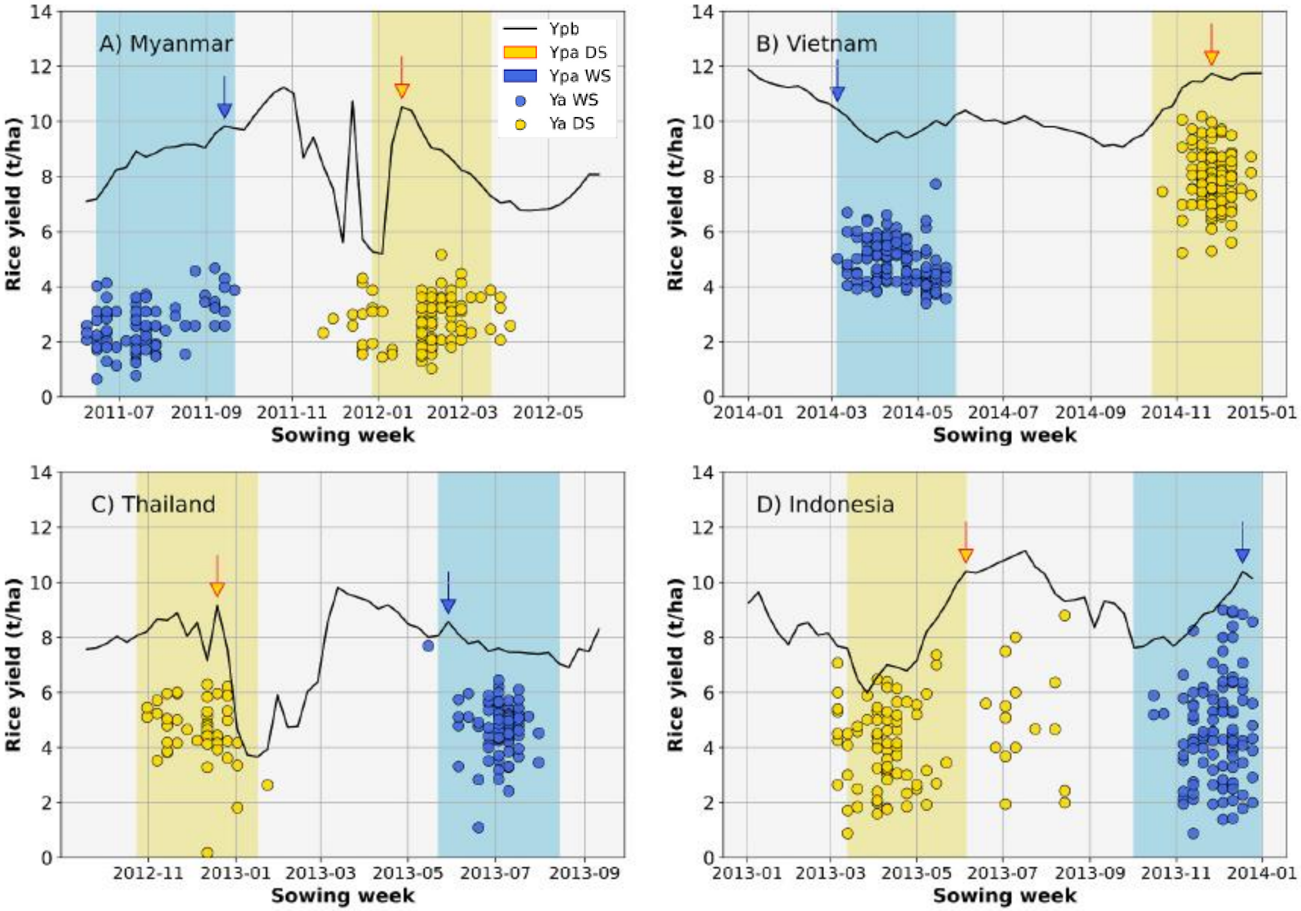
CIMMYT

Simulation of yield levels



Berghuijs et al. (under review)

Simulation of mngt practices



Silva et al. (2022, AgSys)

CIMMYT

Data-driven approaches

Boundary lines
(Fermont et al., 2009)

Field Crops Research 112 (2009) 24–36

Closing the cassava yield gap: An analysis from smallholder farms in East Africa

Stochastic frontiers
(Silva et al., 2017)

Europ. J. Agronomy 82 (2017) 223–241

Explaining rice yields and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling

Machine learning
(Nayak et al., 2022)

Field Crops Research 287 (2022) 108640

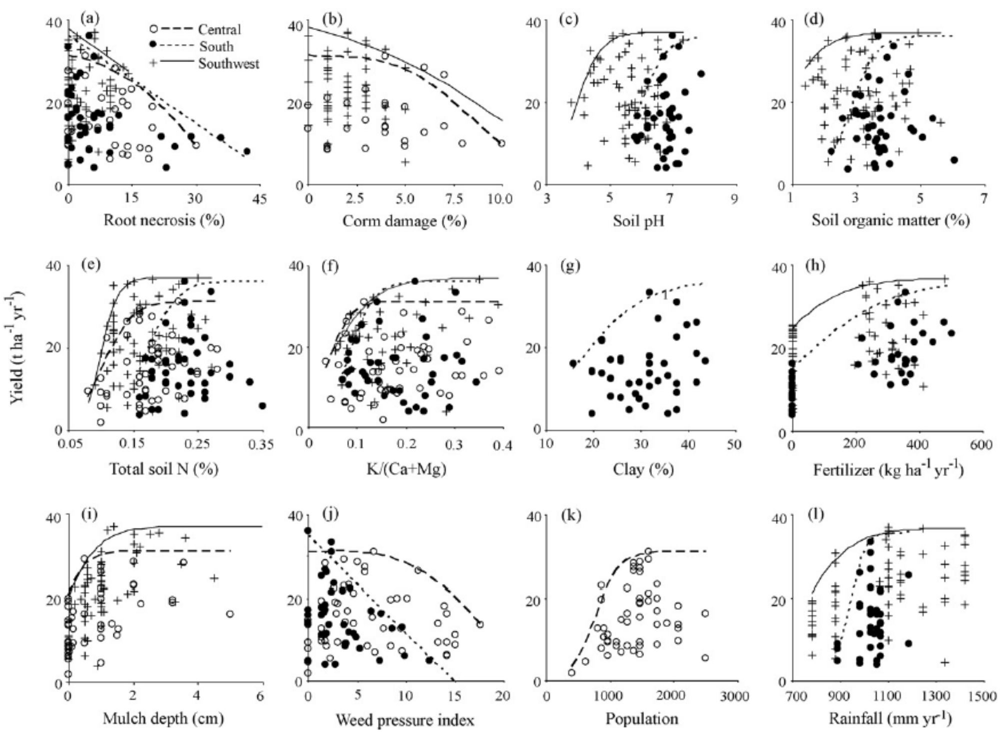
Interpretable machine learning methods to explain on-farm yield variability of high productivity wheat in Northwest India

CIMMYT

Boundary line analysis

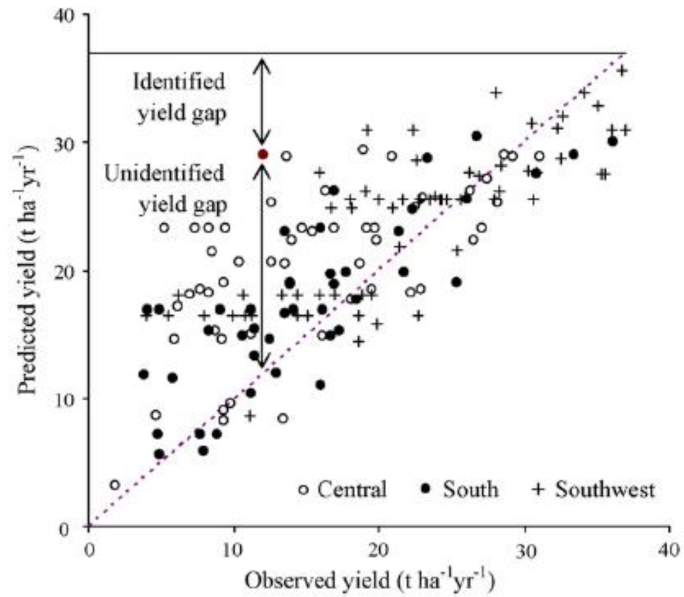
Step 1: Identify boundary line points.

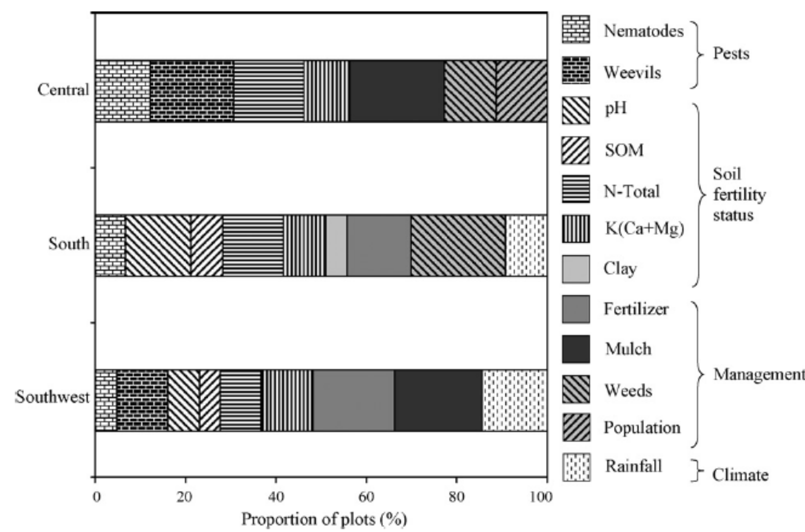
Step 2: Estimate boundary line functions.



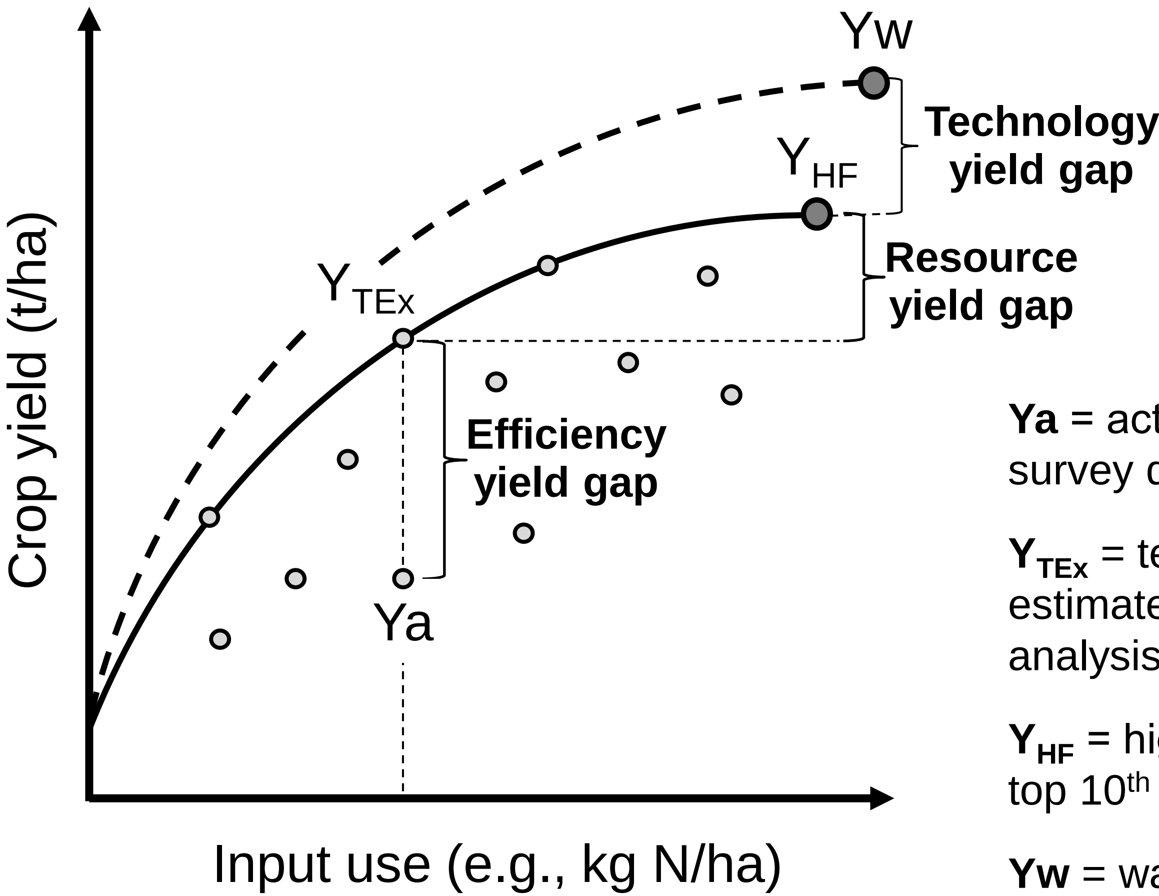
Wairegi et al. (2010, FCR)

Step 3: Yield gap analysis and identification of limiting factors.





Stochastic frontier analysis



Ya = actual farmers' yields from farm survey data

Y_{TEx} = technical efficient yields estimated with stochastic frontier analysis

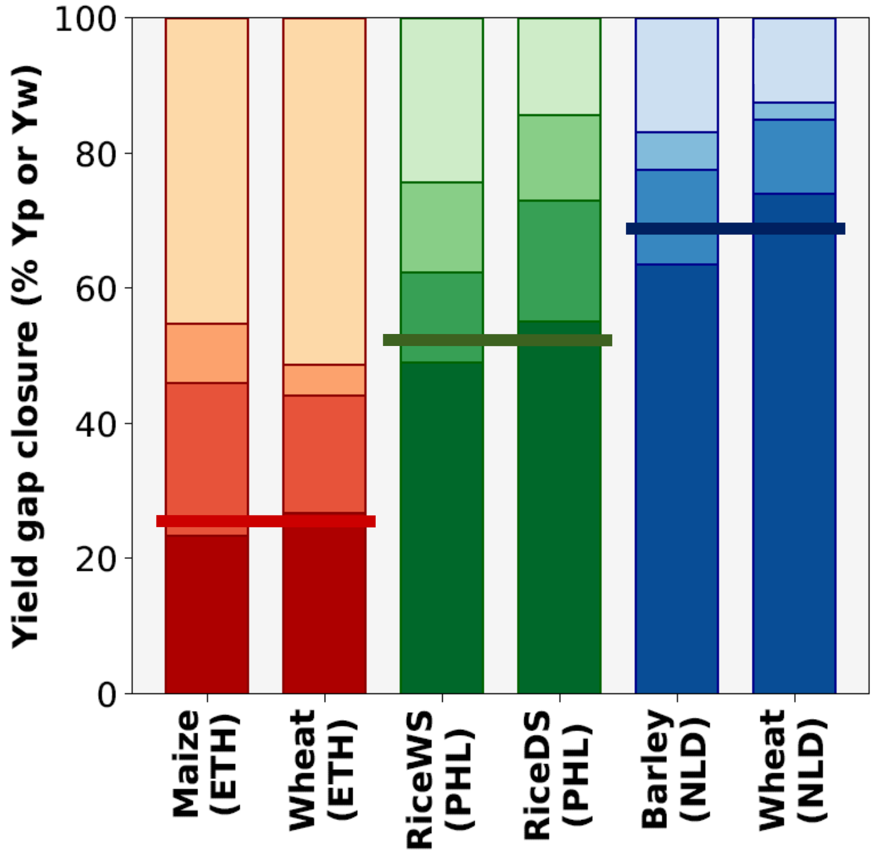
Y_{HF} = highest farmers' yields based on top 10^{th} percentile of Ya

Yw = water-limited potential yield from crop models

Silva et al. (2017, EJA)

CIMMYT

Stochastic frontier analysis



Southern Ethiopia

Large yield gap attributed to technology yield gaps.

Silva et al. (AgSys, 2019)

Central Luzon, Philippines

Medium yield gap due to efficiency, resource and technology yield gaps.

Silva et al. (2017a, EJA)

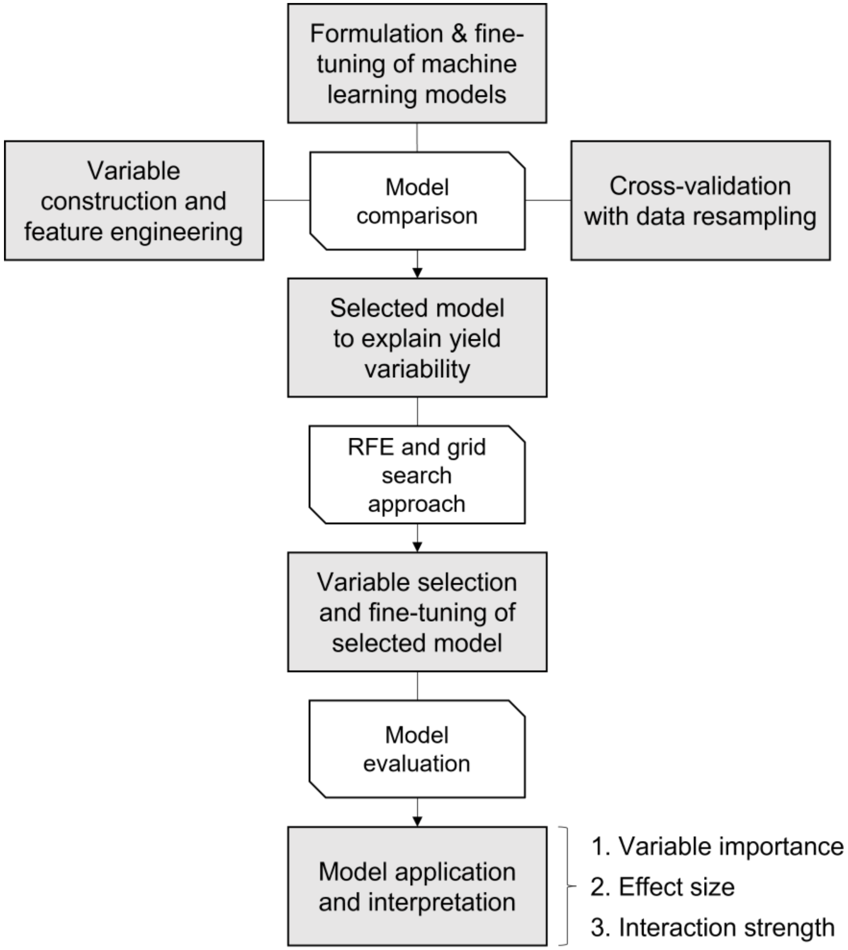
The Netherlands

Small yield gap attributed to efficiency yield gaps.

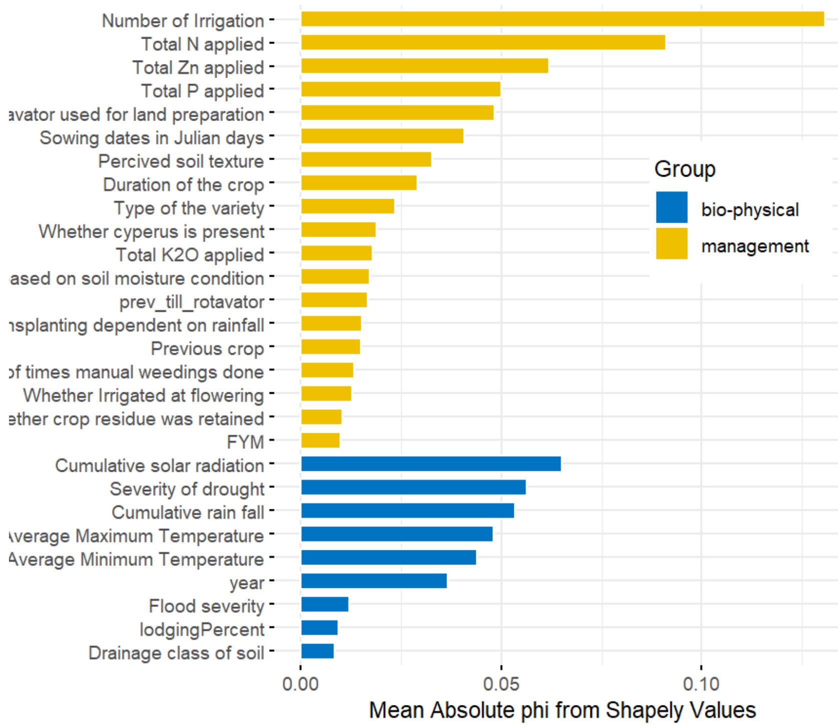
Silva et al. (2017b, AgSys)

Silva et al. (2021, GFS)

Machine learning



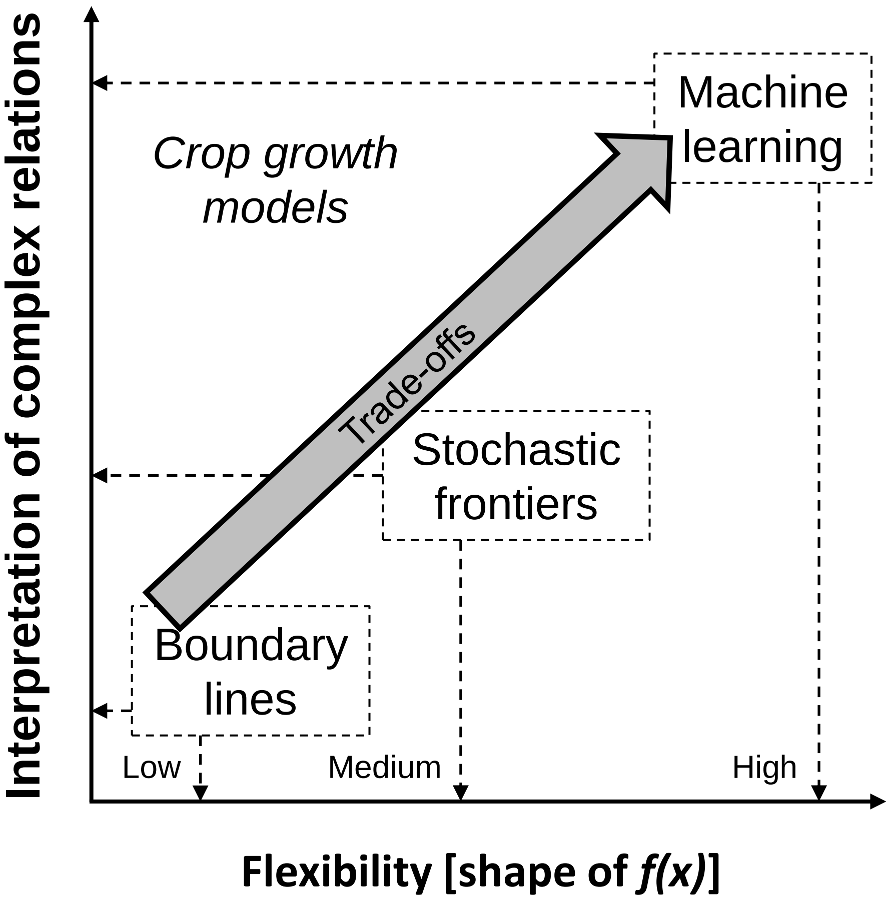
Variable importance/Shapely values



***(!)** Machine learning mostly used for analysis of actual yield variability. Its application for Yg decomposition is being finalized.*

Nayak et al. (2022, FCR)

Summary



Some advantages:

- All types of variables can be used
- Less sensitive to outliers
- Less subjective user decisions
- Non-linear relations and interactions accounted for
- No assumptions on functional forms
- No problem with auto-correlation

Some disadvantages:

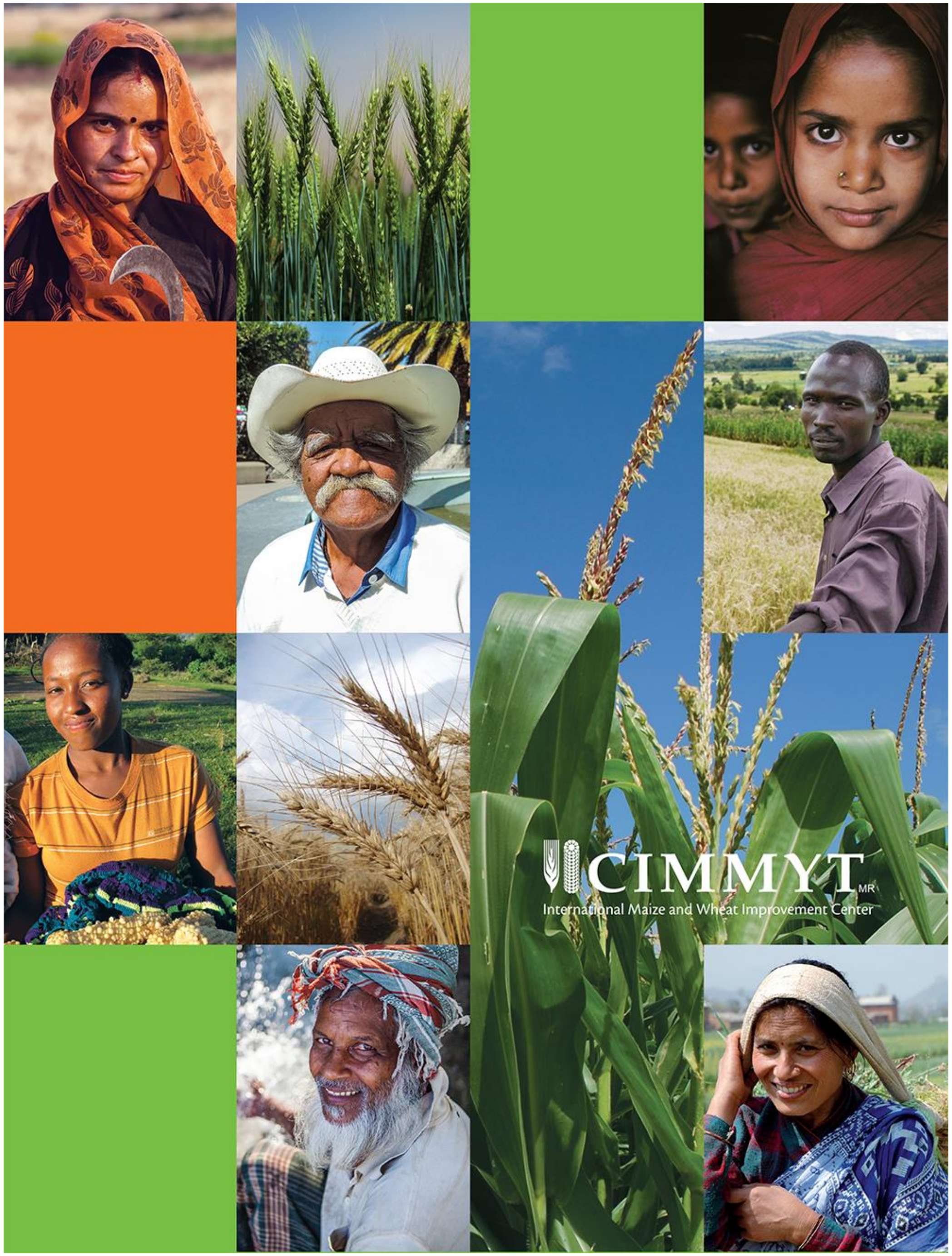
- High sample size
- High computation power
- No interpretability of parameters
- Script-based, Excel won’t work ☺

Strategic (future) applications

Web-based ‘one stop shop’ for yield gap decomposition:

1. Interactive display on constraints to crop productivity in EiA use cases.
2. Measurement and evaluation of agronomic gains (KPIs).
3. Global FAIR database of on-farm yield and management practices.
4. Achieve training materials, data collection tools, publications, etc.
5. Connect and capacitate agronomists and data scientists.

- **Prioritization** of investments and R&D efforts to increase crop productivity in the Global South.
- Inform national and regional **policies** dealing with food security, resource use efficiency, and environmental sustainability.



Thank you for your interest!